

Predictive Learning in Rate-Coded neuronal Networks: A theoretical approach towards classical conditioning

Bernd Porr^a Florentin Wörgötter^a

^a*University of Stirling, Department of Psychology, Stirling FK9 4LA, Scotland,
<http://www.cn.stir.ac.uk>*

Abstract

A novel approach for learning of temporally extended, continuous signals is developed within the framework of rate coded neurons. A new temporal Hebb like learning rule is devised which utilises the predictive capabilities of bandpass filtered signals by using the derivative of the output to modify the weights. The initial development of the weights is calculated analytically applying signal theory and simulation results are shown to demonstrate the performance of this approach. In addition we show that only few units suffice to process multiple inputs with long temporal delays.

Key words: Temporal Hebb; Temporal Sequence Learning; Classical Conditioning

1 Introduction

Classical conditioning takes place on rather long time scales (seconds) and requires two input stimuli, the unconditioned and the conditioned stimulus, which normally come from different sensorial modalities. It can be interpreted as a form of predictive temporal learning where the unconditioned stimulus acts as a predictor of the conditioned stimulus.

Recent theoretical approaches towards temporal learning rest on spiking neurons (3; 5). These models are often highly realistic but usually cannot be treated in a rigorous mathematical way. Also one finds that long time scales and multiple inputs are hard to implement in these models. On the other hand there exists a long tradition of models with linear or so called rate coded neurons which allow an easier treatment (6). Thus, in our approach we have chosen a rate code description as the level of abstraction and present a theoretical framework for predictive (temporal) learning which is able to handle

time-continuous input signals (rate-functions) of arbitrary shape. To this end our “neurons” will act as damped oscillator circuits and a new learning rule is developed which utilises the temporal change (the derivative) of the output to modify the weights. This system has the feature, that very few components are required to cover large time intervals and it allows to use multiple inputs, which are the two critical prerequisites for “classical conditioning”.

In order to implement our algorithm we will introduce neuronal filter circuits which will act as damped oscillators and devise a new learning rule. This rule directly uses the temporal change (the derivative) of the neuronal output activity to modify the weights in the network. In addition, our system has the feature, that very few components are required to cover large time intervals (in contrast to the TD-algorithm) and it allows to use multiple inputs, which are both the critical prerequisites for “classical conditioning”.

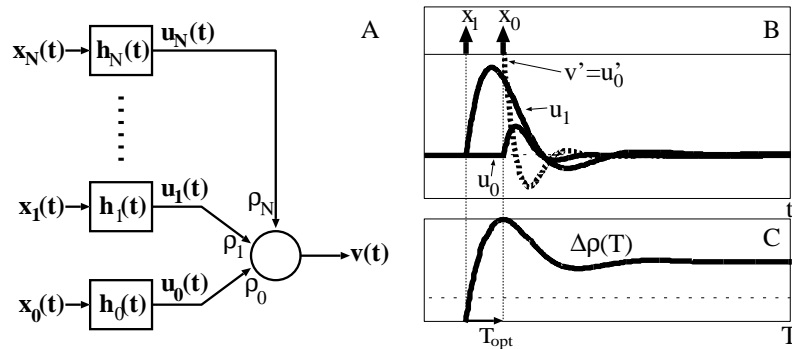


Fig. 1. A) The basic circuit in the time domain. B) shows the inputs x , the impulse responses u for a choice of two different resonators h and the derivative of the output v' . C) shows the initial weight change $\rho_1(T)_{t=0}$. T_{opt} is the optimal time difference between x_1 and x_0 for a maximum change of the weight ρ_1 . The time difference of x_0, x_1 in B is adjusted to T_{opt} , as well. Parameters: $f_0 = 0.1, f_1 = 0.05, Q_0 = Q_1 = 1$.

2 The neuronal circuit

We consider a system of $N + 1$ units h receiving a continuous input signal x and producing a continuous output u . The input units connect with weights ρ to one output unit v (Fig. 1A). All input units are in principle equivalent but we will use h_0 to denote the one unit which transmits the unconditioned stimulus. The output v is then given as:

$$v = \rho_0 u_0 + \sum_{i=1}^N \rho_i u_i \quad \text{with} \quad u_i(t) = x_i(t) * h_i(t) \quad (1)$$

The transfer function h shall be that of a *resonator* with frequency f and quality Q which transforms a δ -pulse input into a damped oscillation (see

Fig. 1B). For later use we will define the transfer function of the resonator in the *Laplace* domain as: $H(s) = [(s + p)(s + p^*)]^{-1}$ with $p = a + ib$, $a = \pi f/Q$ and $b = \sqrt{(2\pi f)^2 - a^2}$.

Learning (viz. weight change) takes place according to a new learning rule changing the weights in each time-step by $\Delta\rho_i$, with:

$$\Delta\rho_i = \mu u_i v' \quad \mu \ll 1 \quad (2)$$

where the weight change depends on the correlation between u_i and the *derivative* of the output v (see Fig. 1B).

3 Results

We state our goal as: After learning, the output unit shall produce a well discernible signal v (e.g., of high amplitude and steeply rising) in response to the earliest occurring conditioned stimulus $x_j, j \geq 1$.

Similar to other approaches (7) we compute the initial development of the weights as soon as learning starts, because this is indicative of the continuation of the learning process. As usual we require that all weight changes occur on a much longer time scale (i.e., very slowly) as compared to the decay of the oscillatory responses u . This allows us to treat the system in a steady state condition, thus, $\Delta\rho \rightarrow 0$ for $\Delta t \rightarrow 0$. In order to calculate this weight change we will now introduce several restrictions: A) The weight of the unconditioned stimulus is set to $\rho_0 = 1$ and kept constant throughout learning. B) We will consider only one unit that can learn, thus, $N = 1$. C) Accordingly we have to deal with only two input functions x_0, x_1 and we define them as (delayed) δ -pulses: $x_0(t) = \delta(t + T), T \geq 0$ and $x_1(t) = \delta(t)$. These restrictions will allow us to develop the theory but can be waived in the end without affecting our basic findings.

Because we assume steady-state, we can rewrite the product in the learning rule (Eq. 2) as a correlation integral between input and output:

$$\Delta\rho_1(T) = \mu \int_0^{\infty} u_1(T + \tau) v'(\tau) d\tau \quad (3)$$

The value of T represents the delay between conditional and unconditional stimulus. In order to assure optimal learning progress we must, therefore, determine that particular value T_{opt} for which a maximal weight change is

observed. Because we are interested in the initial weight change, we can assume $\rho_1(t) = 0$ for $t = 0$. and Eq. 3 turns into:

$$\Delta\rho_1(T)_{t=0} = \mu \int_0^{\infty} u_1(T + \tau)u_0'(\tau)d\tau \quad (4)$$

In order to solve Eq. 4 we have to switch to the LAPLACE domain where $H(s)$ is the transfer function of the resonator as defined earlier. We get:

$$\Delta\rho_1(T)_{t=0} = \mu \frac{1}{2\pi} \int_{-\infty}^{+\infty} H_1(i\omega) [-i\omega e^{T i\omega} H_0(-i\omega)] d\omega \quad (5)$$

We set $P^+ = |p_1|^2 + |p_0|^2$ and $P^- = |p_1|^2 - |p_0|^2$ and get after integration with the method of residues (see curve in Fig. 1 C):

$$\Delta\rho_1(T)_{t=0} = \mu \frac{b_1 P^- \cos(b_1 T) + (a_1 P^+ + 2a_0 |p_1|^2) \sin(b_1 T)}{b_1 (P^+ + 2a_1 a_0 + 2b_1 b_0) (P^+ + 2a_1 a_0 - 2b_1 b_0)} e^{-T a_1} \quad (6)$$

Before showing simulation results we note that the above obtained analytical results can be extended to cover the most general system structure as represented in Fig. 1A. With Eq. 1 transformed in the LAPLACE domain the learning rule turns to:

$$V(s) = \sum_{j=0}^N \rho_j X_j(s) H_j(s), \quad \Delta\rho_i = -\mu \frac{i\omega}{2\pi} \int_{-\infty}^{+\infty} V(-i\omega) U_i(i\omega) d\omega, \quad (7)$$

where the second part of this expression is the general form of Eq. 5. It should be noted that for all $\Delta\rho_i$ this integral can still be evaluated analytically in the same way as in the special case with two resonators discussed above.

In order to validate the approach several simulations of increasing complexity were performed. As the analytical solution treats only the initial learning step we determined the learning behaviour for $t \gg 0$ with a sequence of two δ pulses x_1, x_0 as inputs. Fig. 2A shows that at the initial learning step the output function v still coincides with the unconditioned stimulus response u_0 . After 10 repetitions of the pulse-sequence (Fig. 2 B), the output function has shifted forward to u_1 . This forward shift can be interpreted in the sense of classical conditioning: after learning the output signal v predicts the occurrence of u_0 having been conditioned by u_1 . A sufficiently strong correlation occurs in this setup only for a small temporal difference between the input pulses. To improve on this one can use several resonators h_1, \dots, h_n with different frequencies which will all receive input from the conditioning stimulus x_1 ,

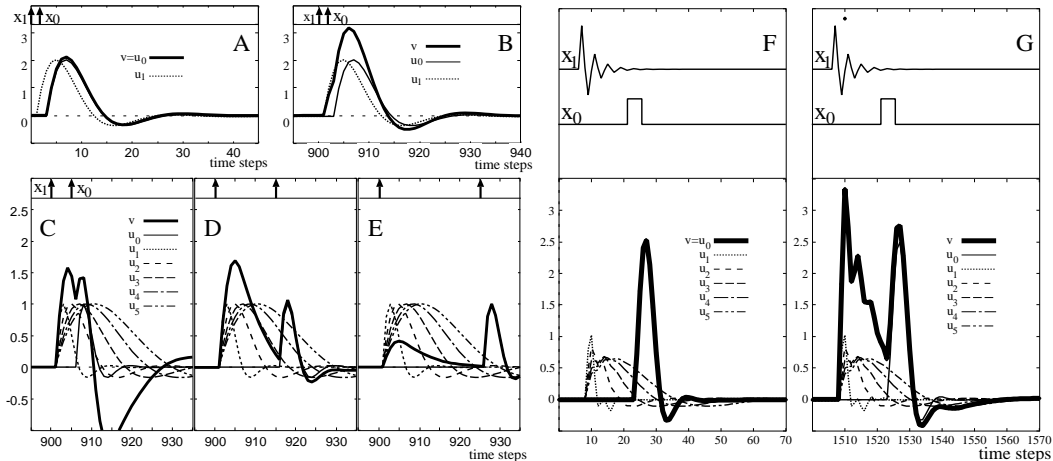


Fig. 2. Simulation results. Arbitrary time steps were used in the simulations. Pulse sequences were repeated every 100 time-steps, the first starting at zero. Resonators were implemented digitally as IIR-filters, which leads to a small onset-delay of 2 time-steps. They had always a value of $a = 0.1\pi$. A,B) are obtained with two identical resonators h_0, h_1 with: $b_{0,1} = 1.09$. The other parameters were $\mu = 0.01$ and $T = 2$. A) Result for $t = 0$, B) for $t = 900$. C-E) are calculated with six resonators, five (curves of u_1, \dots, u_5) receiving input from x_1 and one (curve of u_0) from x_0 . Resonator parameters are: $f_0 = 1.09$, $f_j = \frac{f_0}{j}, j \geq 1$, μ was set to 0.11. Results are shown for $t = 900$. C) $T = 5$. D) $T = 15$, E) $T = 25$. F,G) have the same conditions as D) x_0 was also a square wave but x_1 was sine burst.

defining: $T_0 = T; x_j = x_1, T_j = 0$ for $j \geq 1$ (Fig. 1A and Eq. 7). In Fig. 2 C-E we use five resonators to represent x_1 and show the results after 10 learning steps for increasingly larger intervals T which separate the input pulses. We note that in all cases the output is a superposition of the different resonator outputs and that a well discernible early signal component exists in response to the conditioning stimulus x_1 . This is due to the fact that in this setup the learning process can be seen as a sequential improvement of the rising edge at the onset of the conditioned stimulus. In the next simulation we copied all conditions from the last simulation but replaced the δ -pulses at x_0 and x_1 by temporally extended signals (Fig 2 F,G). This was done to show that the system is able to deal with complex input signals. In an application the signal x_0 (the unconditioned stimulus) could be a light stimulus and the signal x_1 (the conditioned stimulus) a sound event. After learning the system is again able to detect the sine burst as the predictor of the square pulse x_0 .

4 Discussion

From the above results it can be deduced that the system generalises without problems to more generic input combinations (e.g. more than two inputs). An important feature of other algorithms for sequence learning is that after

convergence the output function will approximate an “expectation potential”, which is essentially a rectangular function, starting at x_1 and ending at x_0 (see (2), chapter 10). This is the reason, why our theory was developed using resonators (band-passes), because they allow to compose arbitrary shapes of v , such as expectation potentials. We realize, however, that a pure low-pass characteristic would also suffice to induce temporal learning, which relates our approach to the theory of adaptive filters (4). This helps to understand the predictive properties of our system, because from this context it is known that the derivative of a low-pass filtered signal acts as a predictor of the signal (1). However, adaptive filters rely on the (gradually shifting) self-similarity of a single signal - hence on its auto-correlation properties. Whereas our approach gains its predictive power from the cross-correlation properties between two (or more) signals such as in biological classical conditioning. Therefore, we expect that this approach will also prove useful in a more technologically oriented context.

References

- [1] S.M. Bozic. *Digital and Kalman filtering: an introduction to discrete-time filtering and optimum linear estimation*. E. Arnold, London, 1994.
- [2] P. Dayan and L.F. Abbott. *Theoretical Neuroscience*. MIT Press, Cambridge MA, 2001.
- [3] W. Gerstner, R. Kempter, L. van Hemmen, and H. Wagner. A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383:76–78, 1996.
- [4] S.S. Haykin. *Adaptive filter theory*. Prentice Hall, Englewood Cliffs, 1991.
- [5] R. Kempter, W. Gerstner, and J.L. van Hemmen. Hebbian learning and spiking neurons. *Physical Review E*, 59:4498–4514, 1999.
- [6] Ken D. Miller. Receptive fields and maps in the visual cortex: Models of ocular dominance and orientation columns. In E Donnay, J.L. van Hemmen, and K. Schulten, editors, *Models of Neural Networks III*, pages 55–78. Springer-Verlag, 1996.
- [7] E. Oja. A simplified neuron model as a principal component analyzer. *J Math Biol*, 15(3):267–273, 1982.
- [8] W. Schultz, P. Dayan, and P.R. Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1599, 1997.
- [9] R.S. Sutton and A.G. Barto. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev*, 88(2):135–170, 1981.

Predictive Learning in Rate-Coded neuronal Networks:
A theoretical approach towards classical conditioning

Authors' complete mailing addresses:

Bernd Porr
Department of Psychology
University of Stirling
Scotland FK9 4LA

Phone: +44 01786 467640
+44 01786 466378
Fax: +44 01786 467641
e-mail: bp1@cn.stir.ac.uk

Florentin Woergoetter
Department of Psychology
University of Stirling
Scotland FK9 4LA

Phone: +44 01786 467640
+44 01786 466369
Fax: +44 01786 467641
e-mail: faw1@cn.stir.ac.uk

A tar-file with all TeX-files is stored at
the following address:

http://www.cn.stir.ac.uk/~bp1/cns_paper/